Doble Grado en II+ADE **29/04/ 2019**

Ecometría **Examen Parcial** **Curso 2018-2019**

DNI/NIE : **APELLIDOS:** **Nombre:**

Cada pregunta se resuelve en la hoja de su enunciado, no se pueden responder preguntas distintas en la misma hoja. Las respuestas se deben escribir con tinta azul o negra. Las respuestas deben ser breves pero razonadas. Errores conceptuales importantes pueden afectar a la calificación global del examen.

Teoría(3 puntos). 1) Interpretar geométricamente la regresión lineal múltiple mediante proyectores ortogonales.

2) Suponiendo que el vector de errores sigue una distribución normal del tipo

$$\varepsilon_i \sim \text{ NID}(0, \sigma),$$

demostrar que

$$\frac{1}{\sigma^2}\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}} \sim \chi^2_{n-k}.$$

(Idea: tener en cuenta que $\hat{\boldsymbol{\varepsilon}} = M\boldsymbol{\varepsilon}$ y el valor del rango de M, junto con un ejercicio del tema 1)

3) Definir insesgadez y eficiencia de un estimador, ¿cúales de ellas cumplen los estimadores por mínimos cuadrados de los coeficientes β_i?

4) En regresión lineal múltiple, comparar el sesgo y la varianza de los estimadores de los coeficientes del modelo restringido respecto al completo (modelo restringido: el obtenido al suprimir variables del modelo completo). ¿Qué conclusión se obtiene?

Casi todas las respuestas están más o menos explícitas en los apuntes del curso.

Observación: la respuesta a ¿cúales de ellas cumplen los estimadores por mínimos cuadrados de los coeficientes β_i? en 3), respecto a la eficiencia está en el teorema de Gauss-Markov: propiedad "BLUE"de la matriz de covarianzas, etc.; incluso se podría demostrar que la varianza de cada estimador de los coeficientes (diagonal de la matriz de covarianzas) es la menor posible entre los insesgados lineales, teniendo en cuenta que en la demostración del teorema aparece el producto BB', lo que implica que la diagonal de BB' son los módulos al cuadrado de las filas de B, y por tanto no negativas, de hecho así se demuestra el teorema de Gauss-Markov en regresión simple a partir del de regresión múltiple.

Doble Grado en II+ADE **29/04/ 2019**

Ecometría **Examen Parcial** **Curso 2018-2019**

DNI/NIE : **APELLIDOS:** **Nombre:**

Ejercicio 2(5 puntos). Dados los siguientes datos

y	x_2	x_3
2	2	6
-2	2	4
8	5	2
6	4	3

1) Obtener el plano de regresión de y sobre x_2 y x_3.

2) Dar un valor estimado de la matriz de covarianzas de los estimadores de los coeficientes β_j del modelo. Interpretar el resultado.

3) Obtener el ajuste de regresión de dos modelos restringidos: modelo con variables y,x_3 y modelo constante.

4) Sea el modelo restringido de las variables y, x_3. Relacionar las diferencias entre el modelo restringido y el modelo completo de las estimaciones de los coeficientes β_1 y β_3 con una regresión entre las dos variables explicativas.

Ayuda: $\begin{pmatrix} 4 & 13 & 15 \\ 13 & 49 & 42 \\ 15 & 42 & 65 \end{pmatrix}^{-1} = \begin{pmatrix} 26.3148 & -3.9815 & -3.5 \\ -3.9815 & 0.6481 & 0.5 \\ -3.5 & 0.5 & 0.5 \end{pmatrix}$

1) Con los datos, construimos las matrices

$X = \begin{pmatrix} 1 & 2 & 6 \\ 1 & 2 & 4 \\ 1 & 5 & 2 \\ 1 & 4 & 3 \end{pmatrix}$, $\mathbf{y} = \begin{pmatrix} 2 \\ -2 \\ 8 \\ 6 \end{pmatrix}$. El método de mínimos cuadrados funcionará ya que el rango de X es 3 (p.ej, tomando el menor de las tres primeras filas).

Entonces:

$X'X = \begin{pmatrix} 4 & 13 & 15 \\ 13 & 49 & 42 \\ 15 & 42 & 65 \end{pmatrix}$, $\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix} = (X'X)^{-1}X'\mathbf{y} = \begin{pmatrix} -19.4074 \\ 4.7407 \\ 2 \end{pmatrix}$,

$y = -19.4074 + 4.7407x_2 + 2x_3$, (plano de regresión).

2) La varianza σ del modelo se estima por

$$s^2 = \frac{\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}}{n-k} = 0.2963, \text{ya que } \hat{\boldsymbol{\varepsilon}} = \mathbf{y} - X\hat{\boldsymbol{\beta}},\ n-k=1.$$

Por tanto la matriz de covarianzas se estima por

$$s^2(X'X)^{-1} = \begin{pmatrix} 7.727 & -1.1797 & -1.037 \\ -1.1797 & 0.192 & 0.1481 \\ -1.037 & 0.1481 & 0.1481 \end{pmatrix}.$$

La parte relevante de la matriz, correspondiente a los estimadores de los coeficientes de x_2 y x_3, es

$$\begin{pmatrix} 0.192 & 0.1481 \\ 0.1481 & 0.1481 \end{pmatrix}.$$

Por tanto, tanto los errores en la estimación de los coeficientes, como la correlación entre las dos variables explicativas dan valores pequeños, lo cual es siempre positivo.

3) Modelo (y, x_3):

$$X_1 = \begin{pmatrix} 1 & 6 \\ 1 & 4 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}, X_2 = \begin{pmatrix} 2 \\ 2 \\ 5 \\ 4 \end{pmatrix}, \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_3 \end{pmatrix}_R = (X_1'X_1)^{-1}X_1'\mathbf{y} = \begin{pmatrix} 9.7143 \\ -1.6571 \end{pmatrix}, \; y = 9.7143 - 1.6571x_3.$$

Modelo cte.: $y = \bar{y} = 3.5$.

4) La diferencia mencionada está relacionada con una regresión de x_2 respecto a x_3:

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_3 \end{pmatrix}_R - \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_3 \end{pmatrix} = Q\hat{\beta}_2,$$

$$Q = (X_1'X_1)^{-1}X_1'X_2 = \begin{pmatrix} 6.1429 \\ -0.7714 \end{pmatrix}, \; x_2 = 6.1429 - 0.7714x_3,$$

es decir:

$$\begin{pmatrix} 9.7143 \\ -1.6571 \end{pmatrix} - \begin{pmatrix} -19.4074 \\ 2 \end{pmatrix} = \begin{pmatrix} 29.1216 \\ -3.6570 \end{pmatrix},$$ que se verifica salvo errores de redondeo.

Observación: lo que se ha calculado corresponde al sesgo generado al omitir la variable x_2.

Doble Grado en II+ADE **29/04/ 2019**

Ecometría **Examen Parcial** **Curso 2018-2019**

DNI/NIE : **APELLIDOS:** **Nombre:**

Ejercicio con ordenador (2 puntos) Dados los siguientes datos

y	x_2	x_3
2	2	6
-2	2	4
8	5	2
6	4	3

1) Obtener el plano de regresión de y sobre x_2 y x_3.

2) Dar una estimación de las varianzas de los estimadores de los coeficientes β_j del modelo. Interpretar el resultado.

3) Si consideramos el modelo restringido de las variables x_3 e y, obtener el valor de la diferencia entre el modelo restringido y el modelo completo de las estimaciones de los coeficientes β_1 y β_3.

4) Regresar una variable explicativa adecuadamente respecto a la otra y relacionarlo con lo obtenido en 3).

(se aconseja escribir también en esta hoja las salidas que da el programa de los diversos ajustes por minimos cuadrados, en donde aparezcan los datos pedidos)

En gran medida el ejercicio consiste en verificar mediante ordenador los que se ha hecho en el ejercicio 2.

Representamos sólo la salida de datos relevante para lo que nos piden.

Apartados 1) y 2, modelos completo, salida de datos:

	$\hat{\beta}_j$	s_j
intercepto	-19.4074	2.7923
x_2	4.7407	0.4384
x_3	2.0000	0.3849

1)

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix} = \begin{pmatrix} -19.4074 \\ 4.7407 \\ 2 \end{pmatrix}, \ y = -19.4074 + 4.7407x_2 + 2x_3$$

2) s_j^2 son las estimaciones pedidas: $s_1^2 = 7.7969$, $s_2^2 = 0.192$, $s_3^2 = 0.1481$.

3) Modelo (y, x_3):

	$\hat{\beta}_{j_R}$
intercepto	9.714
x_3	-1.657

$y = 9.714 - 1.657x_3$

Modelo (x_2, x_3):

	$\hat{\beta}_j$
intercepto	6.1429
x_3	-0.7714

$x_2 = 6.1429 - 0.7714x_3$

Como en el ejercicio 2:

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_3 \end{pmatrix}_R - \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_3 \end{pmatrix} = \begin{pmatrix} 29.1214 \\ -3.657 \end{pmatrix} = Q\hat{\beta}_2,$$

con

$$Q = \begin{pmatrix} 6.1429 \\ -0.7714 \end{pmatrix}.$$